Word Frequency Analysis of Dictated Clinical Data: A User-Centered Approach to the Design of a Structured Data Entry Interface

Christopher Kreis [‡]
Paul Gorman, MD [§]
Cornell University Medical College[‡]
Division of Medical Informatics and Outcomes Research, Oregon Health Sciences University[§]

Abstract: The design of a functional interface for direct entry of physical exam data by physicians remains a formidable challenge for developers of clinical information systems. Many developers use a theoretical approach, basing the interface on a model of the structure of the information and of the user-system interaction that is developed with one or more clinical domain expert(s). We explored the use of empirical analysis as a basis for the design of a structured data entry (SDE) interface. A collection of physical examination data from actual trauma patients, dictated by trauma surgeons, was used for the analysis. Using simple parsers written in Visual BASIC™, we used word frequency analysis (WFA) and manual editing to identify the frequencies of unique terms used by physicians in recording 688 HEENT and 712 LUNG physical exams. A second-pass WFA was used to determine associated descriptive terms. A simple SDE interface was created based on the results of these analyses. The interface was then evaluated by assessing the extent to which the HEENT and LUNG segments of similar physical exams could be fully recorded using the empirically-based SDE interface. Using this interface, 68% of 200 trial HEENT exams, and 85% of 200 trial LUNG exams could be fully recorded. The interface was also considered helpful in recording substantial portions of the remainder of the exams. We believe that WFA can be a useful tool for finding empirical basis for SDE design.

INTRODUCTION

Efforts to develop software for structured data entry (SDE) by physicians have been underway for over thirty years,[1] yet the task remains a formidable challenge today.[2] According to the Committee on Improving the Patient Record of the Institute of Medicine, "The single greatest challenge in implementing the CPR [computerized patient record] is to develop a technology...so that [health professionals] can- and will- enter medical and other health care data directly into the computer."[3]

A critical issue in designing an SDE interface is the content and organization of pick lists from which clinicians must choose findings. A given patient's exam requires entry of a small subset of the enormous number of potentially recordable findings. Free text entry allows maximum flexibility, but requires greater keyboarding skill and does not offer the advantages of a constrained, indexable, searchable medical vocabulary. Constrained vocabularies are superior in these respects, but have the disadvantage of reducing the use of clinically rich descriptors. Pick lists, used in many SDE interfaces,[2,4,5,6] can augment the use of a fixed vocabulary, but may be inefficient if intractably large or if not structured in a manner which is intuitive to the user.

In addition, to facilitate use, the interface must be specific to the context, including attributes of the examiner, the patient, and the clinical problem. One standard organization is unlikely to be optimal in all situations. Creating a functional interface thus depends on the ability to predict the subset of findings relevant to the specific context of use,[2,4,5] and organize elements of the SDE interface accordingly. To be usable, the content and organization of elements such as pick lists must conform to the expectations of the clinician in a specific clinical context.

Kushniruk et al., discussing current problems with the CPR, state, "We have not yet learned to represent computer-based medical information in a form that seems intuitive to clinicians."[7] They cite two sets of problems: those associated with the user interface, and those deriving from "conceptual problems that arise when physicians try to map and represent findings with the terms used by the system." Moorman,[9] paraphrasing Kaplan,[16] makes a related point: "SDE has to correspond intuitively to the physician's usual method of working; applications that do not significantly change routine patient care are more likely to be accepted."[8] To address these issues, we explored the use of computerized WFA as a means of defining candidate terms and initial term organization for an SDE interface. We then examined the extent to which the resulting interface could be used to faithfully record physical examination data.

METHODS

Source Data

Source data consisted of a text file containing over 1,000 physical examinations dictated by trauma surgeons at Oregon Health Sciences University as part of routine care. All information identifying specific individual physicians and patients was removed. Dictated data conformed to the standard history and physical examination structure. Transcriptionists routinely typed headings for subsections

1091—8280/97/$5.00

in uppercase, followed by a colon, e.g., "HEENT:"

Programming and Analysis

The dataset was processed in stages using parsers written in Visual BASIC.™ Frequency counts were verified for ten randomly chosen terms, and total word counts were verified, using a commercial word processor. Analysis of frequency count output was performed by manual inspection, assisted by a standard commercial spreadsheet program.

Procedure: Word Frequency Analysis

A first parser identified subsection headings, defined as words followed by colons, and determined their frequencies of occurrence. A second program was used to isolate HEENT and LUNG exam subsections and save these to individual textfiles for separate analysis.

A third program counted the frequencies of occurrences of all words in the HEENT and LUNG exams. Results were transferred to a spreadsheet for manual analysis. Non-relevant words were removed, including prepositions, conjunctions, many verbs, and parts of speech that could not be identified as physical finding descriptors.

Next, associations between words were identified by the frequency of their occurrence in close proximity. For example, the term "clear" occurred frequently in the physical exam data in various contexts. Examining the output of the parser, it could be seen that "clear" most often occurred in proximity to the terms "oropharynx" and "nares." The parser could also be used to identify proximate terms based on a single term, for example, entering "oropharynx" produced a list of commonly used descriptors.

Table 1: Sample WFA Output

Frequencies of Terms Recorded in HEENT Exams

term	*frequency*
heent	666
no	462
pupils	414
clear	396
light	396
reactive	390
equal	374
round	353
oropharynx	337
intact	331
extraocular	316
membranes	302
tympanic	298
movements	234
normocephalic	207

Procedure: Interface Building & Evaluation

A graphical user interface for SDE was created based on WFA data. Terms were selected for the interface for inclusion in the default pick lists based on frequency of occurrence. High frequency terms, mainly describing normal findings, were included on the first page of the on-screen HEENT form. Associated descriptors were likewise included based on frequency of occurrence in proximity. Some commonly encountered abnormal findings were also included on the first screen. Less common terms, those used to describe abnormalities or enhance detail, were included in pop-up boxes. The process was repeated for the LUNG exam: many high frequency words were included in a pop-up menu that described locations of findings. No pop-up box called another pop-up box.

After the entry interfaces were created for HEENT and LUNG exams, 200 HEENT and LUNG exams were randomly selected from the database. The newly created interface was used to attempt to record the findings contained in these exams. An exam was considered to be recordable only if it could be recorded in its full detail, although subjective judgments were necessary in some cases. For example, where "lungs clear to auscultation" had been recorded, it was assumed that "lungs clear to auscultation bilaterally," the term used in the SDE interface, was implied. On the other hand, the term "moderate rales" was considered not recordable; while rales existed in the interface, severity descriptors did not.

RESULTS

Word Frequency Analysis

Using the initial parser, the frequencies of subsection headings were determined for 1,000 exams. The heading "HEENT:" occurred 663 times. Similar headings were "HEAD, EYES, EARS, NOSE, and THROAT:" (19 times), "HEAD:" (4 times), and "HEAD/NECK:" (2 times). Only 688 total such headings were identified out of the 1,000 exams. Many exams did not include separate examination of the head, or a separate subsection so labeled. These exams were excluded from analysis. Using a similar procedure, 712 LUNG exams were identified. Further parsing then identified the frequencies of all words in the exams.

Interface Design

The highest frequency HEENT findings, described from 414 times (pupil findings) down to 18 times (sinus findings), were primarily normal ones. Commonly encountered abnormal findings and descriptors thereof were found to occur from 6 to 18 times. These were the findings included in the HEENT interface. Significantly fewer terms were recorded in the LUNG exam part of the interface. Word frequencies ranged from 571, for "clear" as in "to auscultation," down to about 5 times.

Evaluation

The HEENT and LUNG data entry interfaces were evaluated using new exam data. Using the interfaces, one could fully record 136 of 200 (68%) HEENT exams and 169 of 200 (85%) LUNG exams. Of the exams that could not be fully recorded, the interface still generally aided in recording substantial portions.

DISCUSSION

This study evaluated computerized WFA of clinical data as a means of creating an empirically derived SDE interface for recording medical data. Based on WFA, a simple SDE interface was developed that permitted sample HEENT and LUNG exams to be fully recorded 68% and 85% of the time, respectively. Using more advanced natural language processing techniques, others have examined the completeness of coverage of medical texts by controlled vocabularies such as the UMLS.[17]

In this trauma-based dataset, the content of the LUNG exam was simpler than that of the HEENT exam. Thus, the difference in success of the interface for recording HEENT compared to LUNG exams is not surprising. Exams that could be fully recorded were mainly normal or had common abnormal findings. Of those exams that could not be fully recorded, many contained descriptions of trauma. Trauma descriptions were frequent, but were not uniform. A trauma recording interface is documented in the literature, though its efficacy is not reported.[13] Other exams that could not be fully recorded included subjective descriptions or quotes from the patient. The unique circumstances of trauma and individual descriptions or quotes are examples of the dilemma when attempting to record findings in a standardized or automated fashion without losing information.

The interface is, however, successful at recording most common findings while being far less complex than many paper-based forms for recording physical exams. Additional terms could be added to the current interface without adding too much complexity, though diminishing returns would be expected as less, and less frequently employed terms are added.

Further work is needed to determine the usefulness of this approach for other portions of the exam, such as the abdomen, or for data recorded in other settings, such as the lung when examined by a pulmonary specialist. These exams may have a much greater degree of variability.

Many SDE interfaces are proprietary and information about their design and effectiveness is unavailable. Some well documented interfaces do exist, however.[4,10] Many are for other medical recording tasks, such as radiology results,[11] GI endoscopy,[6] or progress notes.[2] However, the same principles apply and provide interesting insight. Unfortunately, most do not document in detail how they developed the lists of concepts they employ.[2] Of those that do, one representatively employs "a combination of general anatomical and medical knowledge, and specific 'pragmatic knowledge' concerning how doctors - or a particular doctor - prefer to enter information."[5] Another employs "protocol analysis and expert interviews."[12] A third reported that their "Medical Knowledge Base [was] developed by forty physicians (specialists and general practitioners) and revised by five physician-analysts trained in knowledge engineering."[4] Only a few specify having employed chart review.[6] Manual review of large numbers of charts is tedious and time consuming. Automated WFA, however, can make it practical.

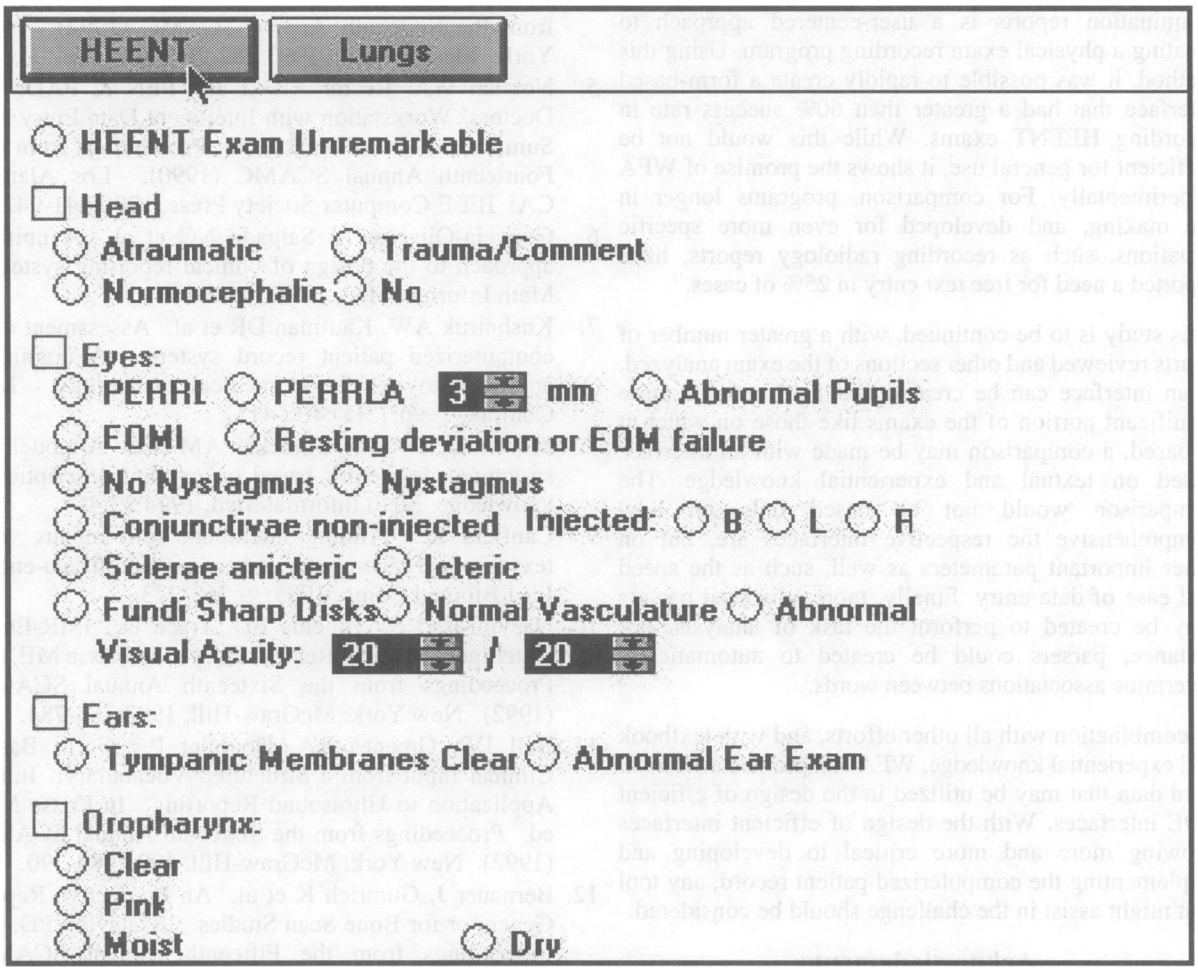
One notable example of how a process similar to WFA might be employed is documented in the Canfield paper.[9] He uses terminology gleaned by computer from a large database of textual echocardiography reports to create the rough initial lists of findings employed in an SDE interface he created. He then tracked term frequencies by computer as interface use was simulated, and reorganized his interface accordingly. Though he states the preference in the literature is for alphabetic organization, he concludes that a combination of alphabetic- and frequency-based organization is optimal.

The advantage of predictability gained by alphabetical sorting is lost as the list length becomes too large for efficient searching. We organized lists according to the frequency with which findings were recorded in our dataset. If frequency is to be employed, WFA provides an empirical basis for determining which terms to include and what their frequencies are. While many clinical texts describe a proposed or ideal structure and language for recording the physical exam, WFA can reveal the language and structure of the physical exam as it is actually recorded.

Use of WFA for interface design, however, also has several potential shortcomings. Using the language of the clinician may be problematic. Some words used by clinicians are not included in standardized medical vocabularies. Standards are of great importance, and the interface that does not employ the standard can have problems, for instance with data exchange. On the other hand, any term that occurs in thousands of medical charts but not in the medical vocabulary, should probably be added.

There are other complications to the technique. It cannot be employed where reports are not dictated. And, while creating an exam based on what clinicians examine may make the interface more intuitive, it may also detract from some of the interface's potential benefits. Clinicians focus on a specific problem and rarely perform a full exam. But in any situation, there are findings that ought to be included, if not as part of addressing the problem at hand, then for prevention and screening (as in the diabetic foot and eye exams. As noted earlier, out of 1,000 exams,

Figure 1 - Portion of Structured Data Entry Screen



examination of the head was not identified in many records. Should a head exam always be recorded in trauma care? The form-based physical exam may have the advantage of prompting the examiner for key findings. Several studies have shown that SDE interfaces do lead to more comprehensive documentation.[11] Excessive promting, however, may reduce physician acceptance or lead to inadvertent inaccurate records. If the form only prompts for findings dictated by WFA, it may leave out findings that texts or experience suggest should be there.

Finally, if the physician becomes accustomed to relying on prompts, he or she may fail to include important elements of the exam not offered in the SDE interface. A low frequency but important finding, such as "alcohol on the breath" (found once in our dataset), can be missed by the recording interface, and might therefore be omitted by the examiner no longer accustomed to check.

There are also potential problems with standardization when the exam based on chart review is so specific to the situation on which it was based. Creation of standards does not involve accommodating varying methods, which is what situation-specific chart review does, but rather settling on one. On the other hand, standardization could be based on chart review. The issue of standards is a critical and complex one in informatics, and is not within the scope of this paper. WFA does demonstrate the great variability among exams in a single domain.

Automated WFA of dictated clinical data cannot be the sole basis for SDE interface development, but it can play an important role. A body of work already exists in this area, and it should not be neglected. Indeed, much of the current focus has moved beyond simply listing findings, but rather creating records based around medical concepts. This would imply, for instance, that instead of placing "conjunctivae injected" on a list, "conjunctivae" would be on a list, and "injected" on another, and the computer could recognize the meaning of their association.[14,15] WFA can serve this model by helping to define terms and their associations.

CONCLUSION

The WFA of many free text dictated physical examination reports is a user-centered approach to creating a physical exam recording program. Using this method, it was possible to rapidly create a form-based interface that had a greater than 60% success rate in recording HEENT exams. While this would not be sufficient for general use, it shows the promise of WFA experimentally. For comparison, programs longer in the making, and developed for even more specific situations, such as recording radiology reports, have reported a need for free text entry in 25% of cases.[11]

This study is to be continued, with a greater number of charts reviewed and other sections of the exam analyzed. If an interface can be created that can record a more significant portion of the exams like those on which it is based, a comparison may be made with an interface based on textual and experiential knowledge. The comparison would not be based only on how comprehensive the respective interfaces are, but on other important parameters as well, such as the speed and ease of data entry. Finally, more advanced parsers may be created to perform the task of analysis. For instance, parsers could be created to automatically determine associations between words.

In combination with all other efforts, and with textbook and experiential knowledge, WFA can provide additional hard data that may be utilized in the design of efficient SDE interfaces. With the design of efficient interfaces growing more and more critical to developing and implementing the computerized patient record, any tool that might assist in the challenge should be considered.

Acknowledgments

Dr. Gorman is supported by a FIRST Award from the National Library of Medicine (LM05663). The authors thank Dr. William Hersh and Mr. Bikram Day for providing the data and helpful comments from their work, supported by NLM Cooperative Agreement LM05879.

References

1 Yoder RD. Preparing medical record data for computer processing. Hospitals, 1966;40(16):75-85.

2. Poon AD, Fagan LM, Shortliffe EH. The PEN-Ivory Project: Exploring user-interface design for the selection of items from large controlled vocabularies of medicine. J Am Med Informat Assoc, 1996;3(2):168-183.
3. Committee on Improving the Patient Record, Institute of Medicine. Dick, RS and Steen EB, eds. The Computer Based Patient Record. Washington, DC: National Academy Press, 1991.
4. Lussier YA, Maksud MS et al. PureMD: A Computerized Patient Record Software for Direct Data Entry by Physicians Using a Keyboard-free Pen-based Portable Computer. In Frisse ME, ed. Proceedings from the Sixteenth Annual SCAMC (1992). New York: McGraw-Hill, 1993:262-264.
5. Nowlan WA, Rector AL et al. PEN & PAD: A Doctors' Workstation with Intelligent Data Entry and Summaries. In Miller RA, ed. Proceedings from the Fourteenth Annual SCAMC (1990). Los Alamos CA: IEEE Computer Society Press, 1990:941-942.
6. Gouveia-Oliveira A, Salgado NC et al. A unified approach to the design of clinical reporting systems. Meth Informat Med, 1994;33:479-487.
7. Kushniruk AW, Kaufman DR et al. Assessment of a computerized patient record system: A cognitive approach to evaluating medical technology. MD Comput, 1996;13(5):406-415.
8. Moorman, PW, van Ginneken AM et al. A model for structured data entry based on explicit descriptional knowledge. Meth Informat Med, 1994;33:455.
9. Canfield K. Priming intelligent split menus with text corpora for computerized patient record data-entry. Int J Biomed Comp, 1995;39:263-273.
10. Naeymi-Rad F, Almeida FD, Trace D. IMR-Entry (Intelligent Medical Record-Entry). In Frisse ME, ed. Proceedings from the Sixteenth Annual SCAMC (1992). New York: McGraw-Hill, 1993:783-784.
11. Bell DS, Greenes RA, Doubilet P. Form Based Clinical Input from a Structured Vocabulary: Initial Application to Ultrasound Reporting. In Frisse ME, ed. Proceedings from the Sixteenth Annual SCAMC (1992). New York: McGraw-Hill, 1993:789-790.
12. Bernauer J, Gumrich K et al. An Interactive Report Generator for Bone Scan Studies. In Clayton PD, ed. Proceedings from the Fifteenth Annual SCAMC (1991). New York: McGraw-Hill, 1992:858.
13. Benoit RG, Cushing BM et al. Direct Physician Entry of Injury Information and Automated Coding via a Graphical User Interface. In Frisse ME, ed. Proceedings from the Sixteenth Annual SCAMC (1992). New York: McGraw-Hill, 1993:787-788.
14. Evans DA, Cimino JJ et al. Toward a medical-concept representation language. J Am Informat Assoc, 1994;1:207-217
15. Rector AL, Glowinski AJ et al. Medical-concept Models and Medical Records: An Approach Based on GALEN and PEN & PAD. J Am Med Informat Assoc, 1995;2:19-35.
16. Kaplan B. The influence of medical values and practices on medical computer applications. In: Anderson JG, Jay SJ, eds. Use and Impact of Computers in Clinical Medicine. New York: Springer-Verlag, 1987:39-50
17. Hersh WR, Campbell EH, Evans DA, Brownlow ND. Empirical automated vocabulary discovery using large text corpora and advanced natural language processing tools. Proc AMIA Annu Fall Symp. 1996:159-63.